

Improvements in Virtual Sensors: Using Spatial Information to Estimate Remote Sensing Spectra

Nikunj C. Oza, Ashok N. Srivastava, and Julienne Stroeve

Abstract—Various instruments are used to create images of the Earth and other objects in the universe in a diverse set of wavelength bands with the aim of understanding natural phenomena. Sometimes these instruments are built in a phased approach, with additional measurement capabilities added in later phases. In other cases, technology may mature to the point that the instrument offers new measurement capabilities that were not planned in the original design of the instrument. In still other cases, high resolution spectral measurements may be too costly to perform on a large sample and therefore lower resolution spectral instruments are used to take the majority of measurements. Many applied science questions that are relevant to the earth science remote sensing community require analysis of enormous amounts of data that were generated by instruments with disparate measurement capabilities. In past work [1], we addressed this problem using Virtual Sensors: a method that uses models trained on spectrally rich (high spectral resolution) data to "fill in" unmeasured spectral channels in spectrally poor (low spectral resolution) data. We demonstrated this method by using models trained on the high spectral resolution Terra MODIS instrument to estimate what the equivalent of the MODIS 1.6 micron channel would be for the NOAA AVHRR/2 instrument. The scientific motivation for the simulation of the 1.6 micron channel is to improve the ability of the AVHRR/2 sensor to detect clouds over snow and ice. This work contains preliminary experiments demonstrating that the use of spatial information can improve our ability to estimate these spectra.

I. INTRODUCTION

THE idea behind Virtual Sensors is that data mining algorithms trained on spectrally-rich (high spectral resolution) data can be used to generate estimates of what those measurements would have been for data that are spectrally-poor (low spectral resolution). This enables us to glean more information from that spectrally-poor data. This is an important problem to solve because spectrally-poor data may be available for longer periods of time than spectrally-rich data. This happens because of improvements in measurement capabilities due to instruments being built in phases, technological improvements, or the need to reduce measurement costs. Many applied science questions that are relevant to the remote sensing community need to be addressed by analyzing very large amounts of data that were generated by instruments with different measurement capabilities.

For example, consider the relationship between the AVHRR/2 (Advanced Very High Resolution Radiometer) and the MODIS (Moderate Resolution Imaging Spectroradiometer)

instruments. AVHRR/2 generates images in only five spectral channels, whereas MODIS generates images in 36 different spectral channels. However, AVHRR/2 data has been available since 1981 whereas MODIS has only been available since 1999. MODIS channels 1, 2, 20, 31, and 32 correspond reasonably well to the five AVHRR/2 channels. We can use data mining methods to model any MODIS channel not available in AVHRR/2 as a function of these five MODIS channels. We can then use the learned model to generate an estimate of what that MODIS channel would have been had it been available in AVHRR/2 given the five actual AVHRR/2 channels as input. If the learned model is of high quality, we can use it to obtain estimates of MODIS channels for years prior to 1999 when MODIS came on-line. We refer to this as a *Virtual Sensor* because it estimates unmeasured spectra. We use Virtual Sensors to generate an estimate of MODIS channel 6 (1.6 microns) for AVHRR/2 because a spectral channel at 1.6 microns is useful for discriminating clouds from snow- and ice-covered surfaces. We chose this task to demonstrate the usefulness of Virtual Sensors. However, in our previous work, we did not take spatial information into account, i.e., predictions were generated for each pixel using only the other channels at that pixel. In this work, we perform some preliminary experiments which demonstrate the extent to which spatial information can improve results. Note that, at the resolution of the images we are working with (1.25km), it is not obvious that spatial information would help.

In the next section, we briefly discuss the scientific motivation for using Virtual Sensors to simulate MODIS channel 6 for the AVHRR/2 instrument. In Section III, we describe Virtual Sensors formally and as a general method going beyond the specific application that we discuss in Section II. In Section III, we discuss the methods we use to incorporate spatial information. In Section IV we discuss our experimental results. Section V concludes the paper and discusses future work.

II. VIRTUAL SENSORS FOR CRYOSPHERE ANALYSIS

Intensification of global warming in recent decades has raised interest in year-to-year and decadal-scale climate variability in the Polar Regions. This is because these regions are believed to be among the most sensitive and vulnerable to climatic changes. The enhanced vulnerability of the Polar Regions is believed to result from several positive feedbacks, including the temperature-albedo-melt feedback and the cloud-radiation feedback. Recent observations of regional anomalies in ice extent, thinning of the margins of the Greenland ice

This work was supported by the NASA Intelligent Systems Intelligent Data Understanding Program.

N. C. Oza and A. N. Srivastava are at the NASA Ames Research Center. J. Stroeve is at the National Snow and Ice Data Center.

sheet, and reduction in the northern hemispheric snow cover, may reflect the effect of these feedbacks. Remote sensing products provide spatially and temporally continuous and consistent information on several polar geophysical variables over nearly three decades. This period is long enough to permit evaluation of how several cryospheric variables change in phase with each other and with the atmosphere and can help to improve our understanding of the processes in the coupled land-ice-ocean-atmosphere climate system. Cloud detection over snow- and ice-covered surfaces is difficult using sensors such as AVHRR/2. This is because of the lack of spectral contrast between clouds and snow in the channels on the earlier AVHRR/2 sensors. Snow and clouds are both highly reflective in the visible wavelengths and often show little contrast in the thermal infrared.

The AVHRR Polar Pathfinder Product (APP) consists of twice daily gridded (at 1.25 and 5km spatial resolution) surface albedo and temperature from 1981 to 2000. A cloud mask accompanies this product but has been found to be inadequate, particularly over the ice sheets [2]. The 1.6 micron channel on the MODIS instrument as well as the AVHRR/3 sensor can significantly improve the ability to detect clouds over snow and ice. Therefore, by developing a virtual sensor to model the MODIS 1.6 micron channel (channel 6) as a function of the AVHRR/2 channels, we can improve the cloud mask in the APP product, and subsequently improve the retrievals of surface temperature and albedo in the product. In doing so we will be able to improve the accuracy in documenting seasonal and inter-annual variations in snow, ice sheet and sea ice conditions since 1981.

III. VIRTUAL SENSORS IN GENERAL

In this section, we discuss Virtual Sensors in general, going beyond the specific application discussed in section II. For purposes of the discussion presented here, we model the data as matrices of time series (following the notation in [3]). The spatiotemporal random function $Z(\mathbf{u}, \lambda, t)$ is modeled as a finite number n of spatially correlated time series with the following representation:

$$\begin{aligned} Z(\mathbf{u}, \lambda, t) &= [Z_{\mathbf{u}}(\lambda, t)] \\ &= [Z_{u_1}(\lambda, t), Z_{u_2}(\lambda, t), \dots, Z_{u_n}(\lambda, t)]^T \end{aligned} \quad (1)$$

In Equation 1, \mathbf{u} represents the spatial coordinate, λ represents the vector of measured wavelength(s), and t represents time. The superscript T indicates the transpose operator. If multiple wavelengths are measured, then each Z_i is actually a matrix, and the function $Z(\mathbf{u}, \lambda, t)$ represents a data cube of size $(n \times \Lambda \times T)$, where these symbols represent the number of spatial locations, the total number of measured wavelengths, and the total number of time samples, respectively. In this notation, the spatial coordinate \mathbf{u} represents the coordinates (or index) of a measurement at a particular location in the field of view. Conceptually, the equation above describes a set of n $(\Lambda \times T)$ matrices. In the event that the spatial coordinate indexes image pixels, it is useful to think of Equation 1 as

describing a time series of data cubes (spectral images) of size $n \times n \times \Lambda$.

Consider a situation where one is given a sensor S_1 which takes k spectral measurements in wavelength bands $\mathbf{B}_1 = \{\lambda_1, \lambda_2, \dots, \lambda_k\}$ at time t_1 . Suppose that we have another sensor S_2 which has a set of spectral measurements taken at time t_2 , $\mathbf{B}_2 = \{\lambda_1, \lambda_2, \dots, \lambda_k, \lambda_{k+1}, \lambda_{k+2}, \dots, \lambda_{k+l}\}$ that partially overlaps the spectral features contained in \mathbf{B}_1 in terms of power in the spectral bands. Thus, \mathbf{B}_1 (or, in general, $\mathbf{B}_1 \cap \mathbf{B}_2$) are the common spectral measurements. Note that these measurements are common only in their power. $\mathbf{B} = \mathbf{B}_2 \setminus \mathbf{B}_1 = \{\lambda_{k+1}, \lambda_{k+2}, \dots, \lambda_{k+l}\}$ represents the measurements available in \mathbf{B}_2 that are not available in \mathbf{B}_1 . We investigate the problem of building an estimator $\Gamma(Z(\mathbf{B}))$ that best approximates the joint distribution $P(Z(\mathbf{B})|Z(\mathbf{B}_1))$, where $Z(\mathbf{B})$ is the data cube for the wavelength bands \mathbf{B} . Thus, we have:

$$\Gamma(Z(\mathbf{B})) \approx P(Z(\mathbf{B})|Z(\mathbf{B}_1)) \quad (2)$$

The value of building an estimator for P is clear particularly in situations where S_1 has been in operation for a much longer period of time than S_2 . S_1 may have fewer spectral channels in which measurements are taken compared to S_2 . However, it may be of scientific value to be able to estimate what the spectral measurements in wavelengths \mathbf{B} would have been if S_1 could have measured them.

The joint distribution given by $P(Z(\mathbf{B})|Z(\mathbf{B}_1))$ contains all the information needed to recover the underlying structure captured by the sensor S_2 . If perfect reconstruction of this joint distribution were possible, we would no longer need sensor S_2 because all the relevant information could be generated from the smaller subset of spectral measurements \mathbf{B}_1 and the estimator Γ . Of course, such estimation is often extremely difficult because there may not be sufficient information in the bands \mathbf{B}_1 to perfectly reconstruct the distribution. Also, in many cases, the joint distribution cannot be modeled properly using parametric representations of the probability distribution since that may require a significant amount of domain knowledge and may be a function of the ground cover, climate, sun position, time of year, and numerous other factors.

In the event that all k wavelength bands in S_1 overlap with a corresponding subset of k bands in S_2 , but S_2 has bands not available in S_1 , the estimation process is more straightforward. When partial overlap occurs between two sensors for a given wavelength, calculations need to be performed to estimate the amount of power that would have been measured in the overlapping bands. This can be done using interpolation methods.

We now outline the procedure for creating a Virtual Sensor. At a minimum, we assume that for sensor S_1 we have measurements $Z_1(\mathbf{B}_1)$ from one image, and for another sensor S_2 we assume that we have another image $Z_2(\mathbf{B}_2)$. The procedure for creating a Virtual Sensor is as follows, assuming that we need to build a predictor for channel b_{k+1} (recall that k is the number of bands in \mathbf{B}_1):

- 1) Find parameters θ that minimize the squared error (or another suitable metric)

$[E[\Gamma(Z_2(\mathbf{B}_1), \theta)] - Z_2(\mathbf{b}_{k+1})]^2$. This is the Virtual Sensor model fitting step.

- 2) Apply Γ to the data from sensor S_1 to generate an estimate of $E[\Gamma(Z_1(\mathbf{b}_{k+1}), \theta)]$. This is the step where the estimation of the unknown spectral contribution occurs.
- 3) Evaluate the results based on science based metrics and other information known about the image.

The procedure described above is standard in the data mining literature. From the remote sensing perspective, it is interesting to see the potentially systematic differences between the performances of the estimator on data from sensors S_1 and S_2 . These will tell us how much the differences between the overlapping bands of the two sensors affect the accuracy of the Virtual Sensor relative to the true sensor.

IV. METHODS

For this paper, we used MultiLayer Perceptrons (MLPs) as our model and the predicted output is always the 1.6 micron channel. However, we experimented with several different sets of inputs. As a baseline, we trained one MLP using as inputs the five MODIS channels identified above just at that pixel. However, we also constructed MLPs that use inputs from each pixel's 5x5 neighborhood. We experimented with two different ways of constructing feature sets from this neighborhood. In the first set of experiments, we performed Principal Components Analysis (PCA) on the five input channels in each 5x5 neighborhood (125 total features) and chose the top five components as inputs to train one MLP and the top ten components as inputs for another MLP. We also trained one MLP to use all 125 features as inputs. In a second set of experiments, we performed a PCA separately within each of the five input channels' 5x5 neighborhoods—this yields 25 features per channel. We trained one MLP to use to the top principal component from each channel and another MLP to use the top two principal components. Unlike the first set of experiments, the second set ensured that each channel contributes to the inputs. We hypothesized that this may yield better results because performing PCA within each channel may better capture texture variations within each channel, which can aid in cloud detection.

V. RESULTS

All the MODIS data used in the analysis were geolocated and gridded to a 1.25 km Equal Area Scalable Earth Grid (EASE-grid) [4] containing the Greenland ice sheet and the surrounding ocean (which is mixture of open water and sea ice). Thirty-four MODIS images from the year 2000 and 2002 were processed. The models were trained on one image from day 140 and tested on thirty-three MODIS images from various days of the years 2000 and 2002 ranging from 140 to 246.¹ This approach maximizes the range of differences in

TABLE I
ACCURACY OF CLOUD DETECTION (OVERALL)

Day/Time	P 10ch	P 10	P 5ch	P 5	P All	5ch
140 1515	0.9316	0.8987	0.9329	0.9049	0.9356	0.9248
140 1830	0.9198	0.9095	0.9246	0.9079	0.9235	0.9207
141 1600	0.9287	0.9026	0.933	0.8998	0.9385	0.9283
141 1735	0.9213	0.9135	0.9291	0.9142	0.9275	0.9237
142 1505	0.9427	0.9222	0.9442	0.9232	0.9494	0.9365
142 1640	0.9264	0.9133	0.9301	0.9101	0.9342	0.9308
142 1820	0.9208	0.9181	0.9235	0.9148	0.929	0.9228
142 1955	0.9007	0.887	0.9038	0.8842	0.9066	0.8995
143 1545	0.9428	0.9221	0.9451	0.9195	0.9517	0.9406
143 1725	0.9298	0.9134	0.9332	0.9092	0.9389	0.9357
143 1900	0.9362	0.922	0.9402	0.9188	0.9449	0.9436
144 1450	0.9502	0.9062	0.9508	0.9066	0.9568	0.9369
144 1630	0.9255	0.9014	0.9306	0.8969	0.938	0.9264
144 1805	0.9268	0.8968	0.9316	0.8908	0.9349	0.9355
145 1355	0.949	0.9186	0.9483	0.9182	0.9547	0.9298
145 1535	0.9461	0.9199	0.9478	0.918	0.9542	0.9417
145 1710	0.9134	0.8901	0.918	0.8874	0.922	0.9159
146 1440	0.9551	0.9241	0.9553	0.926	0.9623	0.9499
146 1615	0.9366	0.9098	0.9387	0.9123	0.9473	0.9381
147 1345	0.9538	0.9286	0.9568	0.9299	0.9608	0.947
148 1610	0.929	0.8663	0.9311	0.8696	0.9371	0.9064
149 1650	0.927	0.8978	0.93	0.8983	0.9389	0.9209
149 1825	0.8872	0.8572	0.8955	0.8549	0.8957	0.8982
151 1950	0.8632	0.7483	0.8792	0.7177	0.8559	0.8723
152 1720	0.8718	0.8479	0.8792	0.8348	0.8907	0.8765
152 1855	0.7779	0.7218	0.7937	0.6993	0.7833	0.7899
153 1625	0.9274	0.8961	0.9295	0.8975	0.9409	0.9286
140 1440	0.9377	0.917	0.9403	0.9167	0.9463	0.9293
153 1550	0.9353	0.9027	0.9381	0.9004	0.9451	0.9355
168 1505	0.9294	0.8905	0.9309	0.8892	0.9401	0.9127
213 1435	0.8969	0.8518	0.8939	0.8714	0.9065	0.8458
226 1405	0.9266	0.8835	0.9196	0.8847	0.9279	0.8828
246 1520	0.8743	0.8484	0.8758	0.844	0.8819	0.8535
MEAN	0.9194	0.8893	0.9229	0.8870	0.9273	0.9146
STD	0.0337	0.0452	0.0310	0.0506	0.0346	0.0338

time of year between the training and test images and allows for analysis of how much prediction loss occurs as a result of this difference.

Table I gives the accuracies of detecting cloud cover (fraction of pixels classified correctly) in the various test images. Table II gives the accuracies in each image just over the Greenland ice sheet and Table III gives the accuracies over water. A threshold of 0.2 was used on the MODIS channel 6 images. That is, a value of at least 0.2 is considered to imply that a cloud is present while a threshold of less than 0.2 implies that a cloud is not present. The threshold of 0.2 was chosen because the MODIS cloud mask team uses this threshold. The first column gives the day (which day of the year out of 365) and time (in 24-hour time) when each image was taken. The first 27 images were taken in the year 2000 while the remaining six images were taken in 2002. The second column ("P 10ch") gives the results of taking the top two principal components in each of the five input channels. The third column ("P 10") gives the results of taking the top ten principal components across all the input channels. The fourth column ("P 5ch") gives the results of using the top principal component in each of the five input channels. The fifth column ("P 5") gives the results of using the top five principal components across the five input channels. The sixth column ("P All") gives the results of using all 125 principal

¹In this paper, we only discuss the results of validating with MODIS images and not AVHRR/2 since the 1.6 micron channel is only available for MODIS images and due to a lack of space.

TABLE II
ACCURACY OF CLOUD DETECTION (ICE SHEET)

Day/Time	P 10ch	P 10	P 5ch	P 5	P All	5ch
140 1515	0.9418	0.9127	0.9423	0.9208	0.9438	0.935
140 1830	0.9125	0.9221	0.9174	0.9215	0.9179	0.8959
141 1600	0.9196	0.9001	0.9256	0.9012	0.9319	0.9124
141 1735	0.895	0.8925	0.9083	0.8952	0.9018	0.8916
142 1505	0.9477	0.9362	0.9494	0.9374	0.9509	0.9439
142 1640	0.9014	0.8868	0.9077	0.8824	0.9118	0.9058
142 1820	0.9052	0.9099	0.9067	0.9077	0.9126	0.8969
142 1955	0.8768	0.8755	0.8791	0.8757	0.8796	0.867
143 1545	0.9413	0.9305	0.9428	0.928	0.948	0.9423
143 1725	0.9289	0.9098	0.9307	0.9059	0.9309	0.9349
143 1900	0.9381	0.928	0.9427	0.9271	0.9425	0.9486
144 1450	0.9584	0.9434	0.96	0.945	0.96	0.9593
144 1630	0.9084	0.9031	0.9147	0.9016	0.9233	0.9114
144 1805	0.9162	0.8988	0.9215	0.8967	0.9249	0.9299
145 1355	0.9674	0.962	0.9677	0.9622	0.9713	0.9672
145 1535	0.9617	0.9521	0.962	0.952	0.9668	0.9628
145 1710	0.9116	0.8977	0.9135	0.8963	0.9196	0.9039
146 1440	0.9695	0.9639	0.9706	0.9655	0.9744	0.9656
146 1615	0.9509	0.9496	0.9538	0.9508	0.9581	0.9533
147 1345	0.9519	0.939	0.9566	0.9438	0.9596	0.9453
148 1610	0.9506	0.9125	0.96	0.917	0.9573	0.9574
149 1650	0.936	0.9257	0.9376	0.9249	0.9453	0.9408
149 1825	0.8528	0.8313	0.8601	0.8307	0.8598	0.8783
151 1950	0.8306	0.7798	0.8539	0.7691	0.8278	0.844
152 1720	0.8501	0.8395	0.8581	0.8237	0.8706	0.8589
152 1855	0.761	0.7076	0.783	0.689	0.759	0.7811
153 1625	0.9392	0.9174	0.9405	0.922	0.9498	0.9414
140 1440	0.9372	0.9283	0.9388	0.9295	0.942	0.9308
153 1550	0.9388	0.9166	0.9428	0.9186	0.9468	0.9392
168 1505	0.9189	0.9029	0.9238	0.899	0.9314	0.914
213 1435	0.8773	0.8513	0.8811	0.8769	0.8919	0.8389
226 1405	0.94	0.8947	0.9294	0.9025	0.9374	0.8938
246 1520	0.8653	0.8638	0.8689	0.8611	0.8726	0.8477
MEAN	0.9197	0.8995	0.9197	0.8994	0.9219	0.9133
STD	0.0443	0.0513	0.0402	0.0548	0.0446	0.0432

TABLE III
ACCURACY OF CLOUD DETECTION (WATER)

Day/Time	P 10ch	P 10	P 5ch	P 5	P All	5ch
140 1515	0.9232	0.8872	0.9251	0.8918	0.9288	0.9124
140 1830	0.9259	0.8988	0.9306	0.8964	0.9282	0.9254
141 1600	0.9373	0.9049	0.9401	0.8984	0.9448	0.9364
141 1735	0.9443	0.9319	0.9473	0.9308	0.95	0.9467
142 1505	0.9381	0.9095	0.9396	0.9103	0.9481	0.9274
142 1640	0.9517	0.9404	0.9529	0.9383	0.957	0.954
142 1820	0.9349	0.9254	0.9386	0.9212	0.9437	0.9394
142 1955	0.9197	0.8962	0.9235	0.8909	0.9281	0.9084
143 1545	0.944	0.9151	0.947	0.9124	0.9548	0.9394
143 1725	0.9306	0.9167	0.9355	0.9122	0.9461	0.9397
143 1900	0.9347	0.9172	0.9382	0.912	0.947	0.9418
144 1450	0.9433	0.8747	0.9429	0.874	0.9541	0.9183
144 1630	0.9432	0.8996	0.9469	0.8921	0.9532	0.9378
144 1805	0.9357	0.8951	0.94	0.8858	0.9433	0.9415
145 1355	0.9293	0.8723	0.9277	0.8711	0.937	0.8868
145 1535	0.9332	0.8934	0.9361	0.8901	0.9439	0.9228
145 1710	0.9152	0.883	0.9222	0.8792	0.9242	0.9185
146 1440	0.9423	0.8888	0.9417	0.8909	0.9516	0.9322
146 1615	0.9222	0.8699	0.9236	0.8737	0.9366	0.9237
147 1345	0.9561	0.9154	0.9571	0.9123	0.9622	0.9334
148 1610	0.9163	0.8393	0.9142	0.8418	0.9253	0.8793
149 1650	0.9149	0.8602	0.9198	0.8625	0.9301	0.8932
149 1825	0.9362	0.8939	0.9457	0.8893	0.9467	0.9438
151 1950	0.9003	0.7124	0.9081	0.659	0.8879	0.9023
152 1720	0.8979	0.858	0.9047	0.8481	0.9149	0.8978
152 1855	0.8014	0.7416	0.8085	0.7137	0.8171	0.817
153 1625	0.9121	0.8687	0.9153	0.8658	0.9293	0.9119
140 1440	0.9382	0.9075	0.9416	0.906	0.95	0.9231
153 1550	0.9321	0.8896	0.9338	0.8833	0.9434	0.9259
168 1505	0.938	0.8804	0.9366	0.8813	0.9472	0.9066
213 1435	0.9134	0.8523	0.9047	0.8667	0.9189	0.8487
226 1405	0.914	0.8731	0.9104	0.8682	0.9189	0.8715
246 1520	0.8816	0.8358	0.8814	0.8299	0.8895	0.8577
MEAN	0.9243	0.8803	0.9267	0.8757	0.9334	0.9141
STD	0.0269	0.0462	0.0264	0.0541	0.0267	0.0306

components and the seventh column ("5ch") gives the results of using just the original five input channels. These results show that spatial information helps, especially if all five input channels are given a chance to present inputs. In particular, the methods that use the principal components in each channel separately consistently outperform the methods that use the top principal components overall. Using all 125 features performs best, but the running time is substantially greater than for all the other methods, which is a significant concern since one would like to use these models to generate predictions for large data archives.

VI. CONCLUSION

In this paper we have presented some preliminary experiments that augment the Virtual Sensors method with spatial information. This continues our previous work, which was aimed at aiding in the discrimination of clouds from snow and ice. This is a challenging problem that is essential to solve in order to map the cryosphere using visible and thermal imagery. Clouds often have spectral reflectances and temperatures similar to snow. Most cloud detection algorithms operationally employ a series of spectral threshold tests to determine if a pixel is clear or cloudy. Having a channel centered around 1.6 microns has significantly improved the ability to discriminate

between clouds and snow using new sensors such as MODIS and AVHRR/3. Unfortunately, a vast amount of data have been collected before these sensors existed that did not have a channel designed to detect clouds over snow and ice-covered surfaces. These data sets have large importance for climate studies since they provide over 20 years worth of observations. Thus, being able to improve the cloud masking abilities of these previous sensors will allow for improved monitoring of several cryospheric variables, such as surface albedo, surface temperature, snow and ice cover.

REFERENCES

- [1] A. Srivastava, N. Oza, and J. Stroeve, "Virtual sensors: Using data mining to efficiently estimate remote sensing spectra," *IEEE Transactions on Geosciences and Remote Sensing, Special Issue on Advances in Techniques for Analysis of Remotely Sensed Data*, vol. 43, no. 3, pp. 590-600, 2005.
- [2] J. Stroeve, "Assessment of greenland albedo variability from the avhrr polar pathfinder data set," *Journal of Geophysical Research*, vol. 33, pp. 989-1034, 2002.
- [3] P. C. Kyriakidis and A. G. Journel, "Geostatistical space-time models: A review," *Mathematical Geology*, vol. 31, no. 6, pp. 651-684, 1999.
- [4] R. Armstrong and M. Brodzik, "Earth-gridded ssm/i data set for cryospheric studies and global change monitoring," in *AI Symposium of COSPAR Scientific Commission A*, Hamburg, Germany, 1995, pp. 115-163.